

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number:

0 577 115 A2

(12)

EUROPEAN PATENT APPLICATION(21) Application number: **93110478.0**(51) Int. Cl.⁵: **H04L 12/40**(22) Date of filing: **30.06.93**(30) Priority: **02.07.92 US 907946**(43) Date of publication of application:
05.01.94 Bulletin 94/01(64) Designated Contracting States:
DE FR GB IT(71) Applicant: **3Com Corporation**
5400 Bayfront Plaza
Santa Clara, CA 95052(US)(72) Inventor: **Reid, Richard S.**
465 Stierlin Avenue No.15
Mountain View, California(US)
Inventor: **Hausman, Richard**
4930 Cherryvale Avenue
Soquel, California(US)
Inventor: **Sherer, Paul William**
1054 Gardenia Way**Sunnyvale, California 94086(US)**Inventor: **Rivers, James P.****730 East Evelyn Avenue No.137****Sunnyvale, California 94086(US)**Inventor: **Zikmund, Cynthia****114 Roble Road****Boulder Creek, California 95006(US)**Inventor: **Connery, Glenn W.****655 South Fair Oaks Boulevard****B301 Sunnyvale, California 94086(US)**Inventor: **Strohl, Niles E.****1885 Birchwood Drive****Tracy, California 95376(US)**(74) Representative: **KUHNEN, WACKER &**
PARTNER
Alois-Steinecker-Strasse 22
D-85354 Freising (DE)(54) **Programmed I/O Ethernet adapter with early interrupts for accelerating data transfer.**

(57) In a Local Area Network (LAN) system, an ethernet adapter exchanges data with a host through programmed I/O (PIO) and FIFO buffers. The receive PIO employs a DMA ring buffer backup so incoming packets can be copied directly into host memory when the PIO FIFO buffer is full. The adapter may be programmed to generate early receive interrupts when only a portion of a packet has been received from the network, so as to decrease latency. The adapter may also be programmed to generate a second early interrupt so that the copying of a large

packet to the host may overlap reception of the packet end. The adapter to begin packet transmission before the packet is completely transferred from the host to the adapter, which further reduces latency. The minimal latency of the adapter allows it to employ receive and transmit FIFO buffers which are small enough to be contained within RAM internal to an Application Specific Integrated Circuit (ASIC) containing the transceiver, ethernet controller, FIFO control circuitry and the host interface as well.

EP 0 577 115 A2

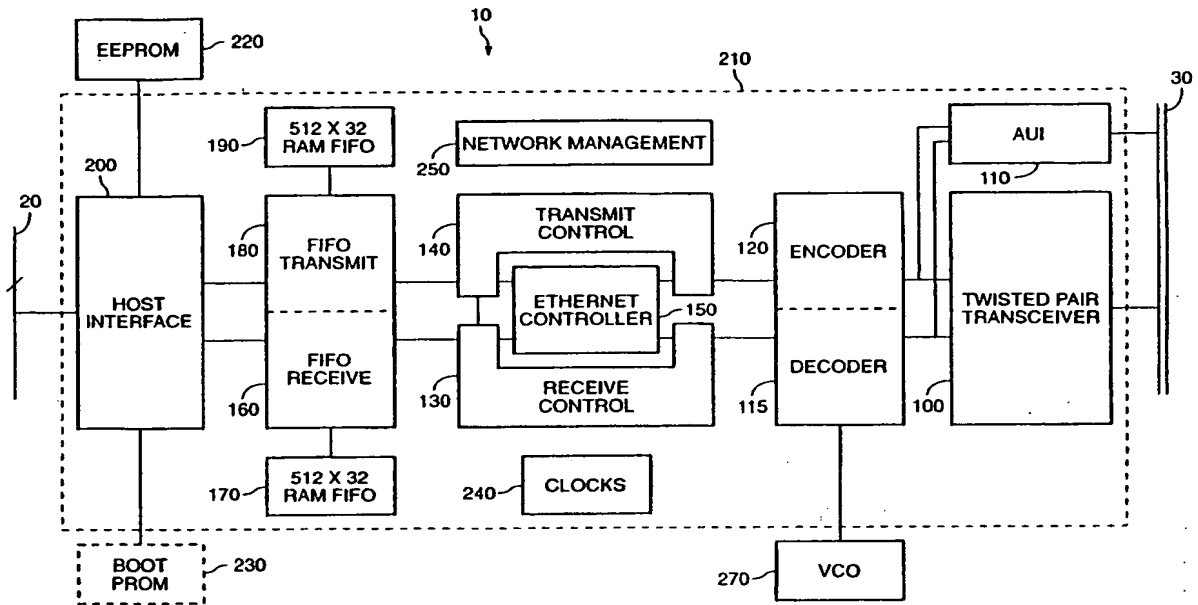


FIG. 1

BACKGROUND OF THE INVENTION

This invention relates to the field of local computer networks, more particularly to an Ethernet adapter providing high throughput for hosts of a network.

Local Area Networks are becoming increasingly common at the office and in industry, where networking enhances productivity by providing improved sharing of information and specialized equipment. Such networks typically consist of an expensive, high capacity server host computer serving a number of relatively less expensive type 286, 386 or 486 Personal Computers as client hosts through which individuals may access the server and specialized equipment. Each host within the network requires an interface apparatus commonly known as an adapter that performs a role intermediate of the host and network for the reception, buffering and transmission of data by the host.

Critical for the usefulness of the PC clients, which comparatively are minimally endowed with speed and memory resources, is an efficient adapter architecture that can allow network communications to proceed in parallel with other computer operations without excessively slowing those other operations. Also critical to the efficiency of the entire network is a need that the adapter have minimal latency in the reception and transmission of data. At the same time, the adapter must be economical to be suitable for accompanying inexpensive computers.

SUMMARY OF THE INVENTION

According to the invention, in a Local Area Network (LAN), a controller in a CSMA/CD (or ethernet) adapter for connecting a host computer node to the network that transfers data to and from the host through programmed I/O (PIO) with first-in-first-out (FIFO) buffers, generates interrupts before complete packets have been received from the network (early receive interrupts), so that reception of the remainder of the packet overlaps with the host computer interrupt latency. The invention reduces overall latency in a CSMA/CD network. As a further aspect of the invention, a second early interrupt may be generated during the reception of large packets so that the copying of the packet to the host may overlap the reception of the final portion of the packet. As a still further aspect of the invention, the adapter is allowed to begin packet transmission before the packet is completely transferred from the host to the adapter, which further reduces latency. The receive PIO employs direct memory access (DMA) ring buffer backup so that incoming packets can be transferred directly into host memory (DMA transferred) when the PIO FIFO

buffer is full.

The minimal latency of the adapter allows the adapter to employ relatively smaller receive and transmit FIFO buffers which can be contained within RAM internal to an Application Specific Integrated circuit (ASIC). Specifically, the ASIC may contain the transceiver, ethernet control circuitry, FIFO control circuitry, FIFO RAM buffers and the host interface in one unit. A further understanding of the nature and advantage of this invention may be realized by reference to the remaining portions of the specification and drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of an ethernet adapter according to the present invention, for use with twisted pair physical media.

Figure 2 is a block diagram of an ethernet adapter according to the present invention, for use with coaxial cable physical media.

Figures 3A and 3B illustrate the structure of transmit data packets and receive data packets, respectively, of a particular embodiment.

Figures 3C, 3D, and 3E illustrate the structure of the adapter status register, the receive status register, and the transmit status register, respectively, of a particular embodiment.

Figure 4 is a flow chart illustrating the transmit process with early transmit interrupts.

Figure 5 is a flow chart illustrating the adapter receive process with early receive interrupts.

Figure 6 is a flow chart illustrating the driver process in response to an early receive interrupt.

DETAILED DESCRIPTION OF PARTICULAR EMBODIMENTS

Referring to Fig. 1, a block diagram is shown of an adapter 10 according to the present invention for communications between a client host computer and an ethernet computer network twisted pair wires. Adapter 10 comprises a number of components coupled between a host computer system bus 20 and communications network twisted pair physical media 30. A twisted pair transceiver 100 is coupled in parallel with an Attachment Unit Interface (AUI) 110 between physical media 30 and both decoder circuitry 115 and encoder circuitry 120. Decoder circuitry 115 is coupled to receive control circuitry 130, and encoder circuitry 120 is coupled to transmit control circuitry 140. Receive control circuitry 130 and transmit control circuitry 140 are coupled to each other and are both coupled to and share ethernet control circuitry 150. Ethernet control circuitry 150 may also include an interrupt timer, explained in more detail below. Receive control circuitry 130 is coupled to receive

First In First Out (FIFO) circuitry 160, which is in turn coupled to a receive (RX) RAM FIFO 170. Similarly, transmit control circuitry 140 is coupled to transmit FIFO circuitry 180, which is in turn coupled to a transmit (TX) RAM FIFO 190. Receive FIFO circuitry 160 and transmit FIFO circuitry 180 are both coupled to a host interface 200 coupled to host system bus 20. Host interface 200 also includes DMA control circuitry.

The transceiver, control circuitry, and RAM discussed thus far are shown in Fig. 1 enclosed by a dotted line to indicate that this circuitry may all be contained within a single Application Specific Integrated circuit (ASIC) 210. ASIC 210 also employs an EEPROM 220 coupled to host interface 200 to provide configuration data such as station address, and it may also optionally employ a boot PROM 230 for automatic configuration. Also contained within and employed throughout ASIC 210 are clocks 240 and network management circuitry 250 to manage various network status signals. An external voltage controlled oscillator (VCO) 270 is also coupled to decoder 115 and encoder 120.

An alternative configuration for an adapter 10' for networks carried by coaxial cable physical media 30' rather than twisted pair physical media 30 is illustrated in Fig. 2. Adapter 10' employs an ASIC 210' that is substantially identical to ASIC 210, except that it uses an external coaxial transceiver 260 rather than internal twisted pair transceiver 100.

Overview of Operation

All data transfer operations between adapter 10 and the host are performed preferably through programmed I/O (PIO), except that a direct memory access (DMA) mode is available as a backup for receive operations. Data is stored by the adapter as double words (4 bytes). As a data packet is received, it is copied into receive FIFO 170. An early receive threshold size is established so that any packet larger than a preselected size triggers the early receive interrupt. If adapter 10 is not provided with or programmed for early receive interrupts, or if the packet is smaller than the early receive threshold size, adapter 10 will wait until the entire packet has been received and then generate an interrupt indicating that a complete packet has been received, that is, a receive complete interrupt, to signal a driver that a complete packet is available for reading. If adapter 10 is provided with or programmed for early interrupts at a particular early receive threshold, an early receive interrupt will be generated once that number of bytes have been received. The driver may then begin reading the data, or for long packets may reprogram the early receive threshold to generate another early receive

interrupt once more of the packet has been received.

As a protection against overflow of the receive FIFO, called receive FIFO overrun, a DMA backup mode may be enabled. If the driver is unable to service receive FIFO 170 adequately, such as if other interrupt handlers consume excessive CPU time, DMA backup will be initiated once receive FIFO 170 has less than a receive FIFO free byte threshold number of remaining available bytes. During DMA mode, data is copied directly from the top of receive FIFO 170 into a DMA ring buffer in the host computer memory.

For transmit operations, all data must be moved into transmit FIFO 190 by the driver through PIO. Typically the driver will copy as much of the packet to the adapter as possible. To minimize latency according to the invention, the adapter may begin transmitting the packet before the complete packet has been copied into transmit FIFO 190. If one or more earlier packets yet remain in transmit FIFO 190, there may be insufficient space for the current packet to be completely copied into transmit FIFO 190. In such a case, the driver will set a threshold to indicate that the transmit function is available, called a TX available threshold, specifying a number of bytes to request an interrupt from adapter 10 when the required number of bytes are free in transmit FIFO 190.

The structure of a specific embodiment of the data packets handled by adapter 10 is illustrated in Figs. 3A and 3B. Fig. 3A illustrates a transmit packet 300 containing a four-byte preamble 305, 14 to 1514 bytes of packet data 310 (which includes destination and source addresses), and padding 315, which is for extending the length of the data string up to a double word boundary. Preamble 305 contains a 1-bit INT field which specifies whether there should be an interrupt upon successful complete transmission, and an 11-bit length field which specifies the number of bytes of actual unpadding packet data plus two bytes of data reserved for use in a future protocol. Prior to transmission of the packet, preamble 305 is stripped off and replaced with a standard 802.3 preamble.

Fig. 3B illustrates a receive packet 320 as stored in both RX FIFO 170 and in the DMA Ring Buffer if used. Packet 320 contains 60 to 1514 bytes of packet data 325 followed by padding 330 to a double word boundary, followed by a postamble 335. Postamble 335 contains two one-bit flags IN and ER, a three-bit code field, and an 11-bit length field which specifies the number of bytes of actual packet data. The IN flag specifies whether the packet is incomplete. The ER flag indicates whether there was an error in the reception, in which case the code field specifies the type of error. These sixteen bits of the postamble duplicate

the RX Status register, described below, and for that reason some embodiments could omit the postamble.

Adapter Control by Driver

Adapter 10 contains numerous registers, some of which may be read by the driver to ascertain the status of adapter 10, others of which may be written to by the driver as commands to control adapter 10, and yet others which are simply used internally by the adapter. In a particular embodiment, these registers are accessed by the driver through a number of eight-word register windows. This method of register access is simply a design choice not critical to the invention, and indeed, many of the commands and registers are not important for an understanding of the invention and need not be described.

One of the primary registers of adapter 10 is the adapter status register, as illustrated in Fig. 3C and indicated generally by reference numeral 350. Its individual bits are briefly described below. Bits marked with an asterisk (*) cause an interrupt when set, unless masked off.

Bit 0	Interrupt Latch-- latched when the adapter raises an interrupt to the host; cleared when acknowledged.
Bit 1	*Adapter Failure
Bit 2	*TX Complete-- the adapter has finished transmitting a packet.
Bit 3	*TX Available-- the number of bytes in the TX FIFO now exceeds the TX Available Threshold.
Bit 4	*RX Complete-- there is a complete packet available in either the RX FIFO or in the host DMA ring buffer.
Bit 5	*RX Early-- more than the RX early threshold, but not all, bytes of the current packet have been received.
Bit 6	*Interrupt Requested-- set when the driver commands the adapter to generate an interrupt.
Bit 7	*Update Statistics-- used for statistics tracking.
Bit 8	(Always zero)
Bit 9	host DMA ring buffer In Use-- indicates the host DMA ring buffer is not empty.
Bit 10	host DMA ring buffer Last Valid-- explained below.
Bit 11	(Reserved)
Bit 12	Command In Progress-- the last command is still being processed

by the adapter.

Bits 13-15 Window Number (0-7)

There are also individual status registers for the receive and transmit FIFOs. The RX status register, as illustrated in Fig. 3D and indicated generally by reference numeral 360, contains two one-bit flags, IC and ER, a three-bit error code, and an 11-bit field, called RX Bytes, containing the receive packet length, which can range from 0 to 1514. References below to RX Bytes should be understood to be to the value contained in the RX Bytes field. Flag IC indicates that the top packet is incomplete or that RX FIFO 170 is empty. Flag ER indicates an error occurred, the type of which is specified in the code field. RX status register 360 is a ripple through FIFO that advances one position when the last byte of a received packet is read from RX FIFO 170. A packet becomes visible in RX status register 360 once either the minimum of 60 bytes have been stored in RX FIFO 170 or the RX Early Threshold has been exceeded. If an error is signalled in the packet before this point, the packet is discarded. If an error occurs after this point, the packet's status will appear in RX status register 360 flagged with an error.

As the packet is received into RX FIFO 170, RX Bytes is incremented. Once the packet has been completely received, the postamble, described above, is written to RX FIFO 170. If the packet is not read from RX FIFO 170 until the incomplete bit is cleared, RX Bytes will show the packet length (assuming there were no errors). As bytes of a packet are read from RX FIFO 170, RX Bytes is decremented. This can be done before the packet has been completely received, in which case RX Bytes shows the number of packet bytes stored in RX FIFO 170. When reading past the end of the packet data, into the postamble, the value RX Bytes is decremented to negative numbers. Reading packet bytes from RX FIFO 170 prior to complete packet reception can be initiated after an initial early receive interrupt through either programming a second early receive interrupt or by simply waiting a period time after the first interrupt. It should also be noted that at any time the driver can issue an RX discard command and the packet will be discarded from RX FIFO 170.

The TX status register, illustrated in Fig. 3E and indicated generally by reference numeral 370, is treated as if stacked in a similar manner to the RX status register. However, it is actually implemented as a pseudo-stack comprising a register of status flags 371 and a five bit counter 372, as opposed to a true stack of status flag registers. The use of such a pseudo-stack is possible because transmission stops upon any transmission error, and does not resume until the error status is popped from the status register. Any sequence of

N transmissions would therefore be a string of N-1 successful transmissions followed by the Nth transmission which may or may not have been successful. Status flags 371 store the status of the Nth transmission, and counter 372 stores N, the number of transmissions in the stack. At a new packet transmission the adapter increments counter 372 and updates flags 371 as necessary. Popping TX status register 370 results in counter 372 being automatically decremented. When the contents of TX status register 370 are accessed, if the count is greater than one, then the regular successful status bits are forced on the bus. If the count is equal to one, then the contents of status flags 371 are applied to the bus, indicating the status of the most recent transmission. If the count is equal to zero, then TX status register 370 is indicated as being empty by applying zeros to the bus.

The flags 371 are a transmission complete flag, a flag specifying whether an interrupt should be generated on successful completion of transmission, and several error flags. Whenever the driver reads TX status register 370 and the TX completed bit is set, the stack is popped, and the next TX status may be read, if any. Popping everything off this stack turns off the TX Complete interrupt bit in adapter status register 350, described above. When the completion of a packet is signalled to the host, the packet has already been discarded from TX FIFO 190. If an error occurred and the packet needs to be retransmitted, it must be copied to TX FIFO 190 again. If the error occurred while the packet was still being copied to the adapter, the host should continue copying the packet to the adapter. When completely copied to the adapter, the packet will be discarded.

Packet Transmission

The basic transmission procedure is performed by the adapter as two independent processes, illustrated by the flow charts of Figs. 4A-B. Fig. 4A illustrates the process performed by the driver when there is packet data to be copied to the adapter, and Fig. 4B illustrates the process performed by the adapter when transmitting packets on the physical medium. As shown in Fig. 4A, the driver first compares at step 400 the number of free bytes in TX FIFO 190 against a TX available threshold. This comparison may be done by polling by the driver, or a TX Available Threshold command may be performed so that the adapter will issue an interrupt when the TX Available Threshold is met. Control remains at this comparison until the TX Available Threshold is met. When the TX Available Threshold is met, control passes to step 405, at which a block of data is copied by the driver into TX FIFO 190. Next, at step 410, it is determined if

more data remains to be copied to the adapter. If not, the process ends until the driver has more packet data to be copied to the adapter. If more data remains, control proceeds to step 415, at which TX FIFO 190 is examined to determine if it is full. If TX FIFO is not full, control passes again to step 405. Otherwise, control returns to step 400. If the packet length specified in packet preamble 305 is less than 60 bytes, the adapter will pad the packet to the 60 byte minimum. Multiple packets can be stored in TX FIFO 190 by employing the process of Fig. 4A.

Fig. 4B illustrates the process by which packets are transmitted from TX FIFO 190 to the physical medium. In step 420 it is determined if the number of bytes in TX FIFO 190 is at least equal to a TX start threshold. The TX start threshold may be varied by the driver to reduce the likelihood of transmit underruns. Control remains at this step until the TX start threshold is met, at which time control passes to step 425. At step 425 the adapter begins to transmit a packet from TX FIFO 190 and continues until the entire packet has been transmitted or an error has occurred, at which point control passes to step 430. At step 430 it is determined if an error was detected during transmission. If so, control passes to step 435 at which it is determined if the error was an underrun error. If the error was an underrun error, then a "bad" CRC is intentionally generated at step 440. If the error was not an underrun, or after a bad CRC is generated, control passes to step 445 where the error status is updated, and next to step 450 where the transmitter is disabled. If at step 430 no error was detected, control passes from there to step 460, at which the transmit status is updated. Next, at step 465, the software protocol (driver) is checked to determine if whether an acknowledgement that packet transmission is complete is required. If not, control returns to start again at the loop of step 420 to await transmission of the next packet. If an acknowledgement is required, at step 470 the driver requests from the adapter an interrupt on successful completion of the next packet transmission, after which control passes to the loop of step 420.

Transmission underruns are generally the result of high interrupt latencies, which are beyond the control of the driver. If a packet underruns, the driver may want to guarantee that the retransmitted packet will not underrun again. This can be done by adjusting the TX start threshold to an amount larger than the packet, so transmission will not begin until the packet is completely copied into the adapter.

A programmable TX Available Threshold is provided by the driver to the adapter to cause the adapter to generate an interrupt when the specified number of bytes become available in TX FIFO 190.

This allows the driver to return and continue copying the data into the adapter at a later time when some of the data in TX FIFO 190 has been transmitted. If TX Available is used with a specified amount less than the size of the next packet to be transmitted, and only a portion of the packet is copied into TX FIFO 190, the driver may want to adjust the early TX threshold to larger than that portion of the packet, to prevent an underrun. This decision may be based upon whether the size of the packet portion in TX FIFO 190 is larger than the amount that can be transmitted during the expected interrupt latency.

Packet Reception by Adapter

Illustrated by the flow chart of Fig. 5 is the basic receive procedure performed by the adapter. The process begins in step 500, where the adapter waits for the beginning of a new packet to be received, after which control passes to step 505. In step 505 the packet 802.3 preamble is received and then discarded by the controller. Next, in step 510 the next byte of the packet is received and stored in RX FIFO 170, after which RX Bytes is incremented in step 515. As mentioned above, the partially received packet, and its associated status values, will not become visible to the driver until RX Bytes reaches the sixty byte minimum or exceeds the RX Early Threshold.

Subsequently, in step 520, if DMA backup is enabled, RX Free, the number of free bytes remaining in RX FIFO 170, is compared to the DMA threshold. If insufficient bytes remain in free, then control passes to step 525, where the DMA process is begun (described in more detail below). If sufficient bytes remain, control passes to step 530.

At step 530 it is determined whether the entire packet has been received. If so, execution passes to step 535, where the RX status register is adjusted accordingly, after which an RX Complete interrupt is generated in step 540 before returning to step 500. If the entire packet had not been received at step 520, execution passes to step 545, where the Early RX interrupt enablement is examined. If not enabled, control returns to step 510 to receive more of the packet. If Early RX interrupts are enabled, then control passes from step 545 to step 550, at which RX Bytes is compared to the Early RX threshold. If RX Bytes is less than the Early RX threshold, then control returns to step 510. Otherwise, control passes to step 555 at which an Early RX interrupt is generated to signal the driver that it may begin copying the packet to the host. After step 555, control returns to step 510.

DMA Backup During Receive

DMA backup of PIO copying of data from RX FIFO 170 to the host is advantageous because the host CPU may become overly delayed by other interrupts and unable to service RX FIFO 170 quickly enough. The DMA backup employs a DMA Ring Buffer consisting of a contiguous block of memory between 256 and 16K bytes in length, located in the host memory and accessed through system bus 20. The DMA circuitry, contained within host interface 200, is set up once during initialization, if DMA backup is selected, to access a DMA channel to the DMA ring buffer in host memory. The DMA channel is programmed to transfer data into the receive ring in a manner causing it to automatically wrap around at the end of the DMA ring buffer space. Packets stored in the DMA ring buffer have the same structure as those in RX FIFO 170.

If DMA mode is initiated, the DMA controller will begin copying bytes from the top of RX FIFO 170 into the DMA ring buffer, while receive circuitry 130 may be continuing to add data to the bottom of RX FIFO 170. The DMA controller preferably copies bytes into the DMA ring buffer at a slightly faster rate than receive circuitry 130 adds bytes to RX FIFO 170. Three registers are maintained to provide necessary information to the driver: the host DMA ring buffer Read index, the host DMA ring buffer Write index, and the host DMA ring buffer Last index. The host DMA ring buffer Read index specifies the address of the next byte to be read from the DMA ring buffer by the driver. This register is only writable by the driver and must be maintained by it. The host DMA ring buffer Write index specifies the address to which the next byte will be written by the DMA controller. If the host DMA ring buffer Read index equals the host DMA ring buffer Write index, the DMA ring buffer is empty. A full condition is reached when the host DMA ring buffer Write index comes sufficiently close to the host DMA ring buffer Read index that the next DMA transfer (typically a burst of four or eight bytes) would cause the two to collide on the network. The host DMA ring buffer Last index specifies the address of the last receive packet postamble copied into the DMA ring buffer. Normally, no more than one complete packet would be present in the DMA ring buffer, although the host DMA ring buffer Last index, together with the length specified in the pointed-to postamble, can be used to trace through a series of packets in the DMA ring buffer. It should be noted that operations involving DMA ring buffer addresses should be performed modulo the DMA ring buffer size, so as to properly wrap around at the DMA ring buffer space limits.

Copying of Received Data by Driver

When the driver responds to a Receive Complete interrupt or an Early Receive interrupt, it first checks the host DMA ring buffer In Use bit of the adapter status register. If the DMA ring buffer is in use, the driver should first empty the DMA ring buffer before disabling DMA and then servicing RX FIFO 170.

Otherwise, if the DMA ring buffer is not in use and the driver is responding to a Receive Complete interrupt for a valid packet, it simply begins copying the packet to the host. A packet with an error is discarded.

If the driver is responding to an Early Receive interrupt, it follows the procedure illustrated by the flow chart of Fig. 6. In step 600 the Early Receive interrupt is acknowledged, which disables further Early Receive interrupts. Next, in step 610, it is determined if this is the first Early Receive interrupt for this packet. This can be done by comparing the adapter's Early Receive threshold to an Early Receive threshold value maintained by the driver. If the latter is smaller, then the adapter had been programmed for a second Early Receive interrupt for this packet, in which the packet should be ready for copying to the host, and control passes to step 620. After the packet is copied to the host in step 620, control proceeds to step 630, where the Early Receive threshold is reset to the value maintained by the driver. Next, Early Receive interrupts are re-enabled in step 640, and the driver exits.

If in step 610 it was indicated that this was the first Early Receive interrupt for this packet, control passes to step 650. If the host computer protocol allows early packet indications, then in step 650 the driver compares RX Bytes to the early lookahead size of the protocol to determine if the Early Receive threshold properly accounts for the CPU's interrupt latency. If the two compared values differ by a significant amount, the Early Receive threshold is adjusted accordingly, and at this point the early lookahead portion of the packet is copied to a dedicated early lookahead buffer for the protocol. The interrupt timer incorporated into ethernet control circuitry 150 may instead be used to determine whether the Early Receive threshold should be adjusted (and may be used to determine a need for similar adjustments to the Early Transmit interrupt). Next, in step 660, RX Bytes is compared to the packet length specified in the RX Status register. If the packet has a substantial number of bytes remaining to be received, such that the driver would empty RX FIFO 170 significantly before the last portion of the packet was completely received, in step 670 it is determined to program the adapter for a second Early Receive interrupt, and control

passes to step 680. In step 680 the adapter is programmed for an Early Receive threshold equal to the length of the packet less the number of bytes that would be received during the interrupt latency. After step 680, execution passes to step 640, described above.

The invention has now been explained with reference to specific embodiments. Other embodiments will be apparent to those of ordinary skill in the art. It is therefore not intended that this invention be limited, except as indicated by the appended claims.

Claims

1. A data communications adapter apparatus for coupling a host computer to a computer network employing communications media, the adapter comprising:
 - a) a transceiver coupled to receive and transmit data over the media;
 - b) a transmit data buffer;
 - c) data transmit control circuitry coupled to said transceiver, to said transmit data buffer, and to said host computer, for generating a packet transmit signal causing said transceiver to begin transmitting data from said transmit data buffer over said communications media;
 - d) a receive data buffer; and
 - e) data receive control circuitry coupled to said transceiver, to said receive data buffer, and to said host computer, for storing data received by said transceiver in said receive data buffer, and for generating a receive interrupt signalling to said host computer that data has been received by said transceiver, wherein said data receive control circuitry is operative to generate a receive interrupt once said transceiver has received over said communications media a predetermined number of bytes of a data packet less than all of said data packet.
2. The adapter of claim 1, further comprising:
 - a) ethernet control circuitry; and
 - b) host interface circuitry, wherein said ethernet control circuitry, said host interface circuitry, said data receive control circuitry, said data transmit control circuitry, said receive data buffer and said transmit data buffer are all contained in a single Application Specific Integrated Circuit (ASIC).
3. The adapter of claim 1, wherein said data receive control circuitry is programmable.

4. The adapter of claim 1, wherein said data transmit control circuitry is programmable to generate said packet transmit signal when said transmit data buffer contains a predetermined number of bytes of a data packet less than all of said data packet. 5
5. A method of transferring a packet of data from a computer network communications media through an adapter to a host computer, said method comprising the steps of: 10
 - a) receiving from said communications media through a transceiver and storing in an adapter receive buffer a predetermined first receive threshold number of bytes of said packet; 15
 - b) thereupon generating a first early receive interrupt from said adapter to said host computer; and
 - c) thereafter receiving from said communications media through said transceiver and storing in said adapter receive buffer a remainder of said packet. 20
6. The method of claim 5, wherein said host computer employs a driver allowing for early indications and having an early lookahead size, and wherein step of receiving a predetermined first receive threshold number of bytes comprises receiving a number of bytes substantially equal to said early lookahead size. 25 30
7. The method of claim 5, wherein said host computer employs a driver allowing for early indications and having an early lookahead size, wherein said adapter and said host computer together have an interrupt latency time, said method further comprising after said step of receiving a predetermined first receive threshold number of bytes, the step of continuing to receive from said communications media through said transceiver and store in an adapter receive buffer bytes of said packet, wherein said step of receiving a predetermined first receive threshold number of bytes comprises receiving a predetermined first receive threshold number of bytes substantially equal to said early lookahead size less a predetermined data receive rate times said interrupt latency time. 35 40 45
8. The method of claim 7, further comprising, after the step of generating said first interrupt, the steps of 50
 - a) evaluating the number of bytes stored in said adapter receive buffer against said early lookahead size; and 55
 - b) adjusting said receive threshold if said evaluating step does not indicate substantial

equality.

9. The method of claim 5, wherein said packet comprises a preamble specifying a length of said packet in bytes, said method further comprising after said step of generating a first early receive interrupt and before said step of receiving a remainder of said packet, the steps of:
 - a) adjusting said receive threshold according to said length of said packet;
 - b) continuing to receive from said communications media through said transceiver and store in an adapter receive buffer bytes of said packet; and
 - c) thereafter generating a second early receive interrupt from said adapter to said host computer, prior to complete reception of said packet.
10. The method of claim 5, wherein said adapter receive buffer has a predetermined size, said method further comprising the steps of:
 - a) if additional bytes of a data packet are received from said communications media through said transceiver and stored in said adapter receive buffer while said adapter receive buffer contains allotted bytes of at least one packet not yet completely transferred to said host computer, such that the total number of allotted bytes in said adapter receive buffer is equal to said size of said adapter receive buffer less a predetermined adapter receive buffer free byte threshold, initiating a direct memory access (DMA) mode; and
 - b) receiving from said communications media through said transceiver and storing in a host computer receive buffer through DMA additional bytes of data.
11. A method of transferring a packet of data from a host computer through an adapter to a computer network communications media, said method comprising the steps of:
 - a) transferring from said host computer and storing in an adapter transmit buffer an early transmit threshold number of bytes of said packet;
 - b) thereupon initiating transmission of said packet through said transceiver from said adapter transmit buffer to said computer network communications media; and
 - c) receiving from said host computer and storing in said adapter receive buffer a remainder of said packet, at least partially in parallel with transmitting said packet through said transceiver from said adapter

transmit buffer to said computer network communications media.

12. A method of transferring through an adapter packets of data between a host computer and a computer network communications media, said method comprising the steps of:
- a) receiving from said communications media through said transceiver and storing in an adapter receive buffer a predetermined first receive threshold number of bytes of a first packet;
 - b) thereupon generating a first early receive interrupt from said adapter to said host computer;
 - c) receiving from said communications media through said transceiver and storing in said adapter receive buffer a remainder of said first packet;
 - d) transferring from said host computer and storing in an adapter transmit buffer an early transmit threshold number of bytes of a second packet;
 - e) thereupon initiating transmission of said second packet through said transceiver from said adapter transmit buffer to said computer network communications media;
 - f) receiving from said host computer and storing in said adapter receive buffer a remainder of said second packet, at least partially in parallel with transmitting said second packet through said transceiver from said adapter transmit buffer to said computer network communications media.
13. The method of claim 12, wherein each packet comprises a preamble specifying a length of said each packet in bytes, said method further comprising after said step of generating a first early receive interrupt and before said step of receiving a remainder of said first packet, the steps of:
- a) adjusting said receive threshold according to said length of said first packet;
 - b) continuing to receive from said communications media through said transceiver and store in an adapter receive buffer bytes of said first packet; and
 - c) generating a second early receive interrupt from said adapter to said host computer prior to complete reception of said first packet.
14. The method of claim 13, wherein said adapter receive buffer has a predetermined size, said method further comprising the steps of:
- a) if additional bytes of a data packet are received from said communications media

through said transceiver and stored in said adapter receive buffer while said adapter receive buffer contains allotted bytes of at least one packet not yet completely transferred to said host computer, such that the total number of allotted bytes in said adapter receive buffer is equal to said size of said adapter receive buffer less a predetermined adapter receive buffer free byte threshold, initiating a direct memory access (DMA) mode; and

- b) receiving additional bytes of data from said communications media through said transceiver and storing through DMA said additional bytes in a host computer receive buffer.

15. An electronic register stack structure for providing, for each of a plurality of actions, a condition represented by a predetermined number of condition flags, wherein said actions may have, upon completion, one of a first condition and at least one alternate condition, wherein a condition of one action at a time is currently accessible from said register stack structure, wherein the currently accessible condition may be removed from said register stack structure so as to make accessible a condition of a remaining action by popping said register stack structure, wherein storing conditions of said actions is halted upon an action having one of said at least one alternate condition, said register stack structure comprising:

- a) a condition register comprising said predetermined number of condition flags;
- b) a counter register;
- c) means for storing a condition for a new action to said register stack structure by modifying contents of said condition register as necessary to indicate said condition of said new action and incrementing said counter register, wherein said counter indicates whether said currently accessible condition is to be said condition of said new action or a condition of a remaining action;
- d) means for providing said currently accessible condition by
 - i) if said counter indicates that said currently accessible condition is to be said condition of said new action, providing contents of said condition register, and
 - ii) if said counter indicates that said currently accessible condition is to be a condition of a remaining action, providing said first condition; and
- e) means for popping said register stack structure by decrementing said counter.

16. The register stack structure of claim 15, wherein said first condition is a successful condition and each said at least one alternate condition is an error condition.

5

17. The register stack structure of claim 15, wherein said means for providing said currently accessible condition further provides a zero condition when said counter indicates that all action conditions have been popped from said register stack structure.

10

15

20

25

30

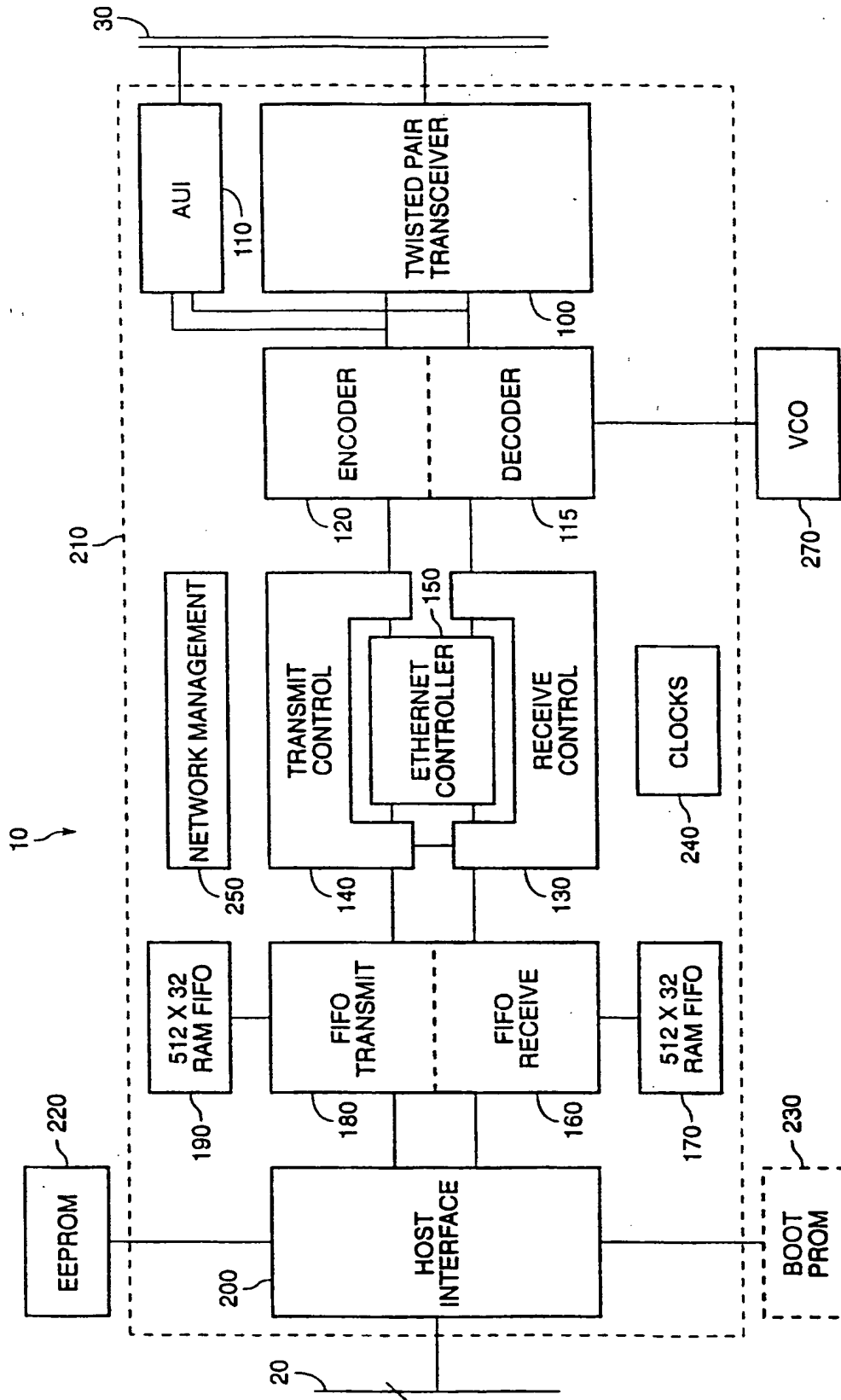
35

40

45

50

55



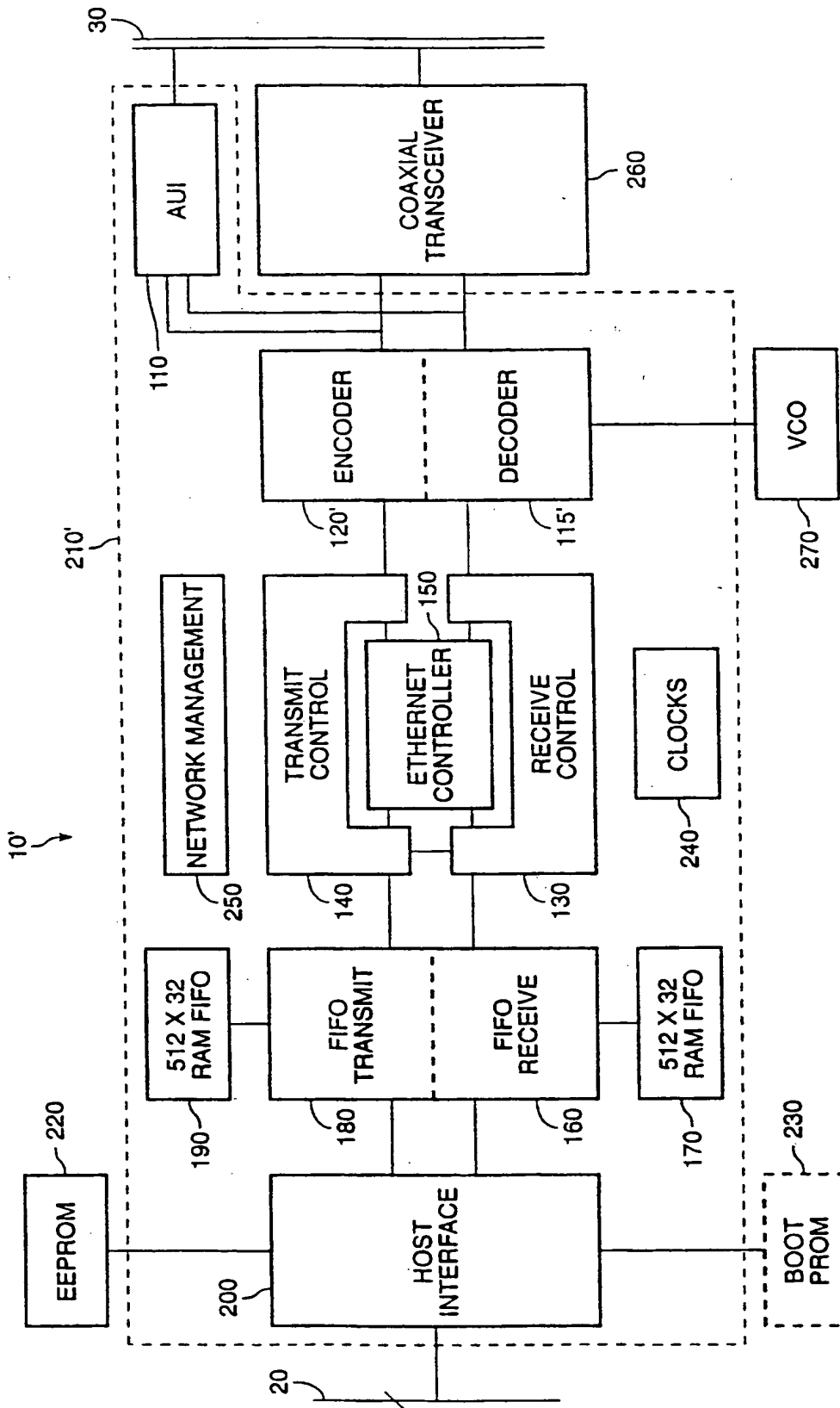


FIG. 2

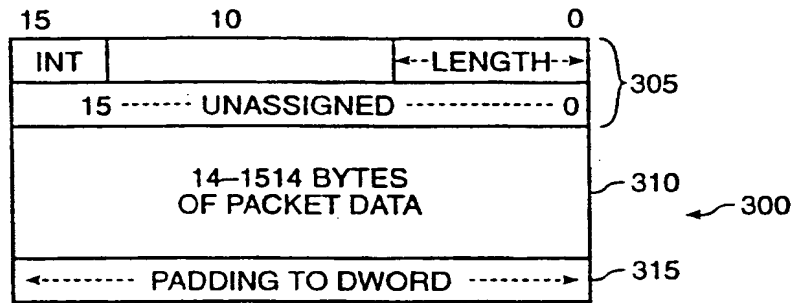


FIG. 3A

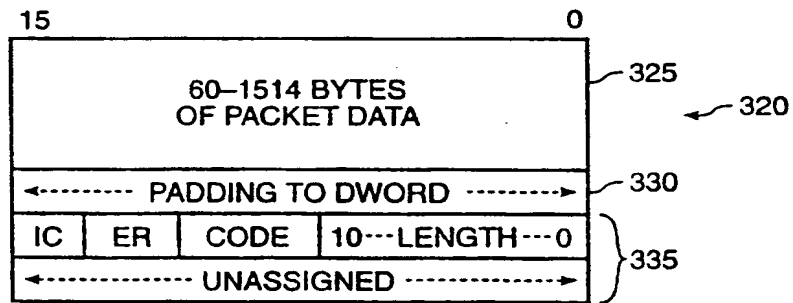


FIG. 3B

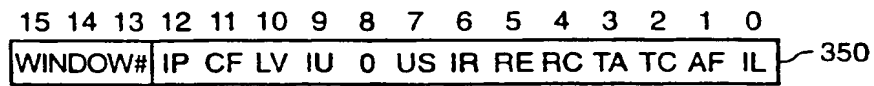


FIG. 3C

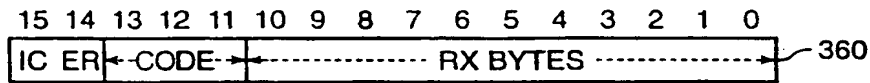


FIG. 3D

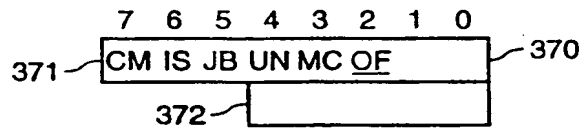


FIG. 3E

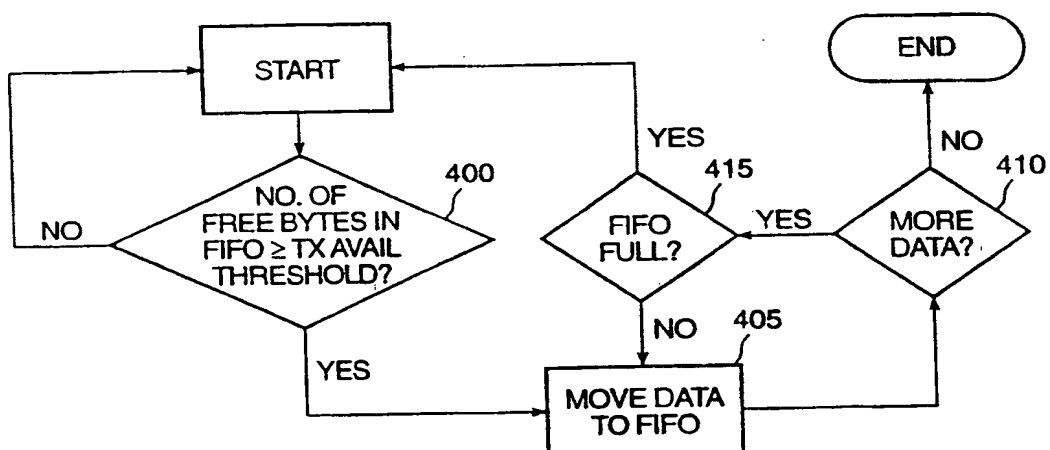


FIG. 4A

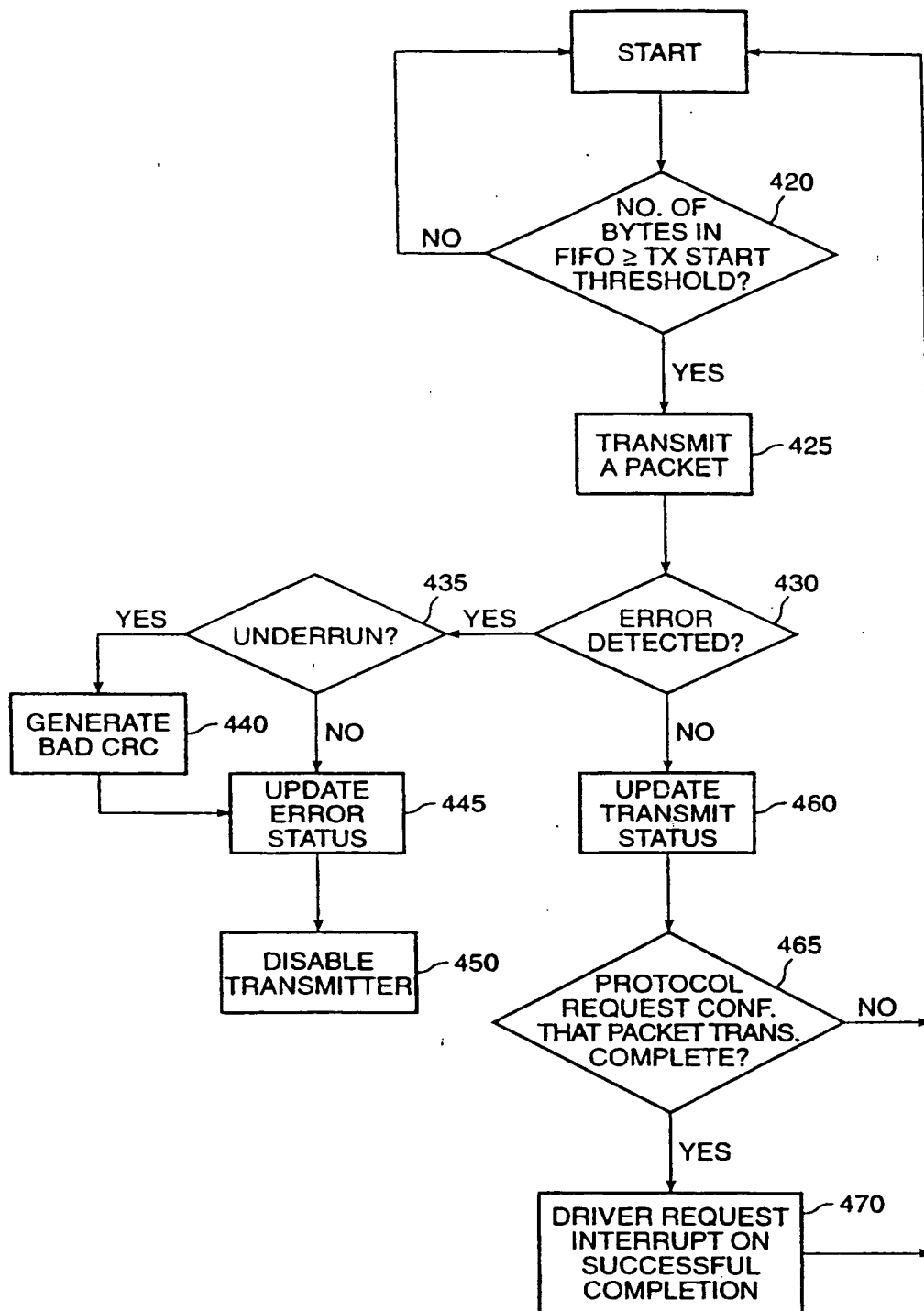


FIG. 4B

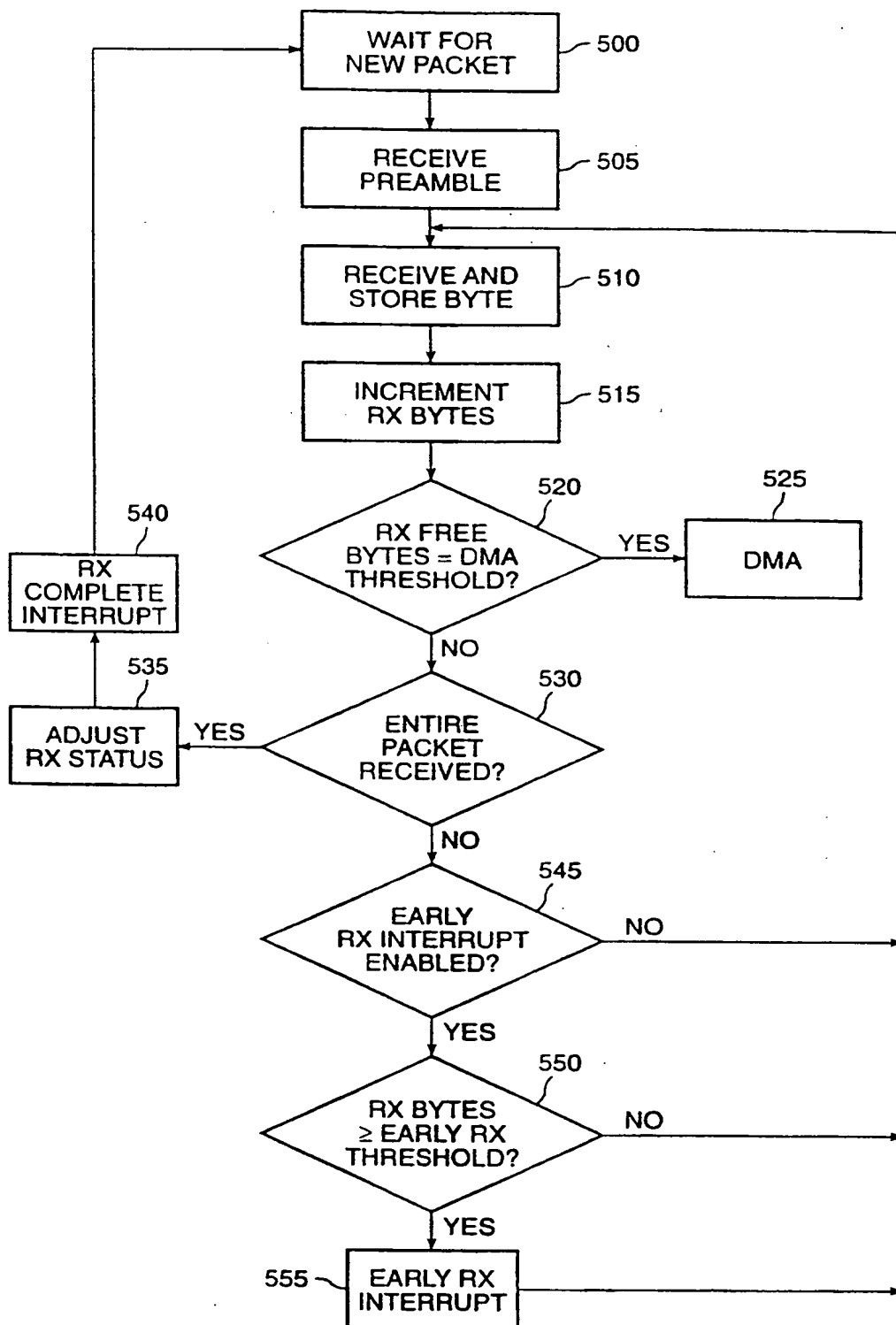


FIG. 5

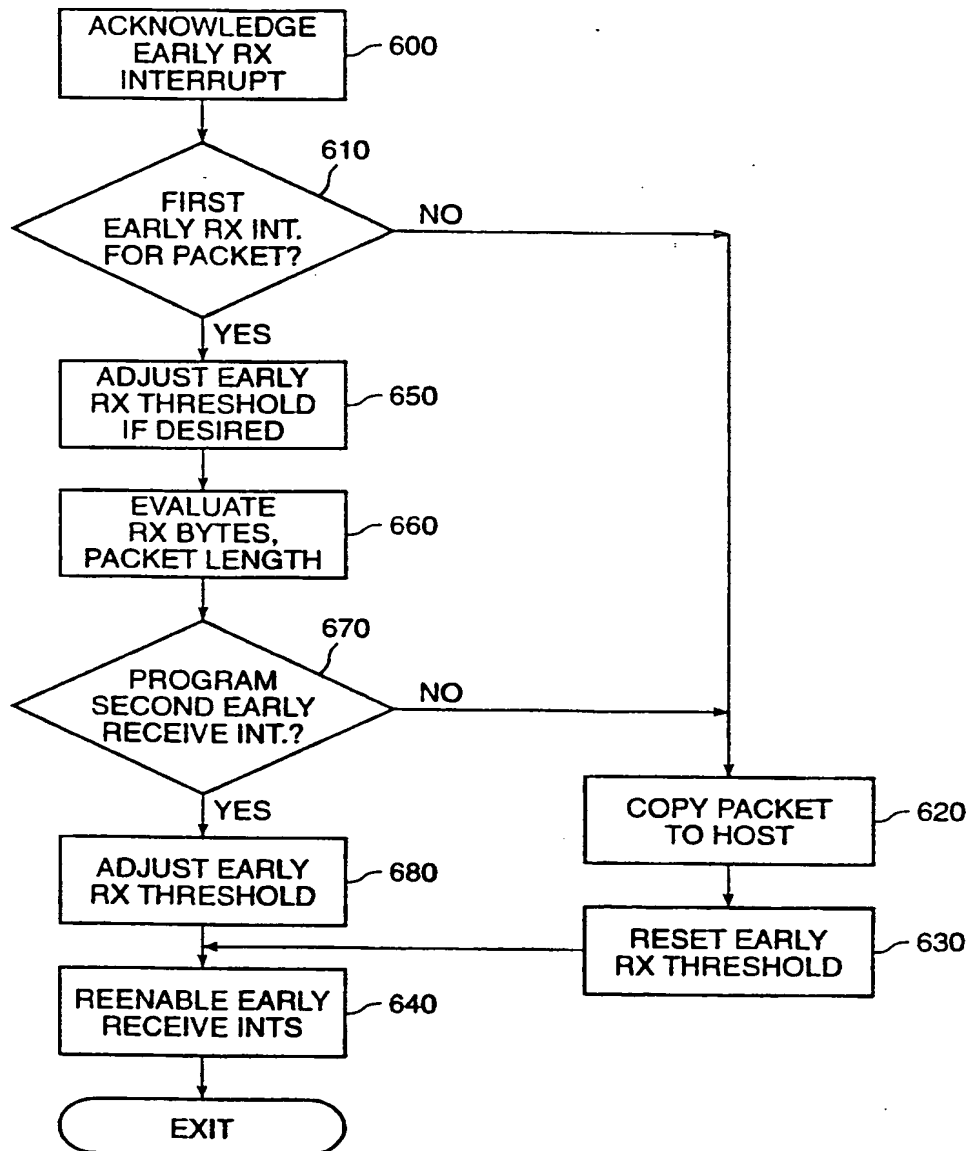


FIG. 6



Europäisches Patentamt
European Patent Office
Office européen des brevets



Publication number: **0 577 115 A3**

12

EUROPEAN PATENT APPLICATION

21 Application number: 93110478.0

51 Int. Cl.⁶: **H04L 12/40, H04L 29/06,
G06F 13/12**

22 Date of filing: 30.06.93

30 Priority: 02.07.92 US 907946

43 Date of publication of application:
05.01.94 Bulletin 94/01

84 Designated Contracting States:
DE FR GB IT

88 Date of deferred publication of the search report:
19.07.95 Bulletin 95/29

71 Applicant: 3Com Corporation
5400 Bayfront Plaza
Santa Clara, CA 95052-8145 (US)

72 Inventor: Reid, Richard S.
465 Stierlin Avenue No.15
Mountain View,
California (US)
Inventor: Hausman, Richard
4930 Cherryvale Avenue
Soquel,
California (US)
Inventor: Sherer, Paul William

1054 Gardenia Way
Sunnyvale,
California 94086 (US)
Inventor: Rivers, James P.
730 East Evelyn Avenue No.137
Sunnyvale,
California 94086 (US)
Inventor: Zikmund, Cynthia
114 Roble Road
Boulder Creek,
California 95006 (US)
Inventor: Connery, Glenn W.
655 South Fair Oaks Boulevard
B301 Sunnyvale,
California 94086 (US)
Inventor: Strohl, Niles E.
1885 Birchwood Drive
Tracy,
California 95376 (US)

74 Representative: KUHNEN, WACKER &
PARTNER
Alois-Steinecker-Strasse 22
D-85354 Freising (DE)

54 Programmed I/O Ethernet adapter with early interrupts for accelerating data transfer.

57 In a Local Area Network (LAN) system, an ethernet adapter exchanges data with a host through programmed I/O (PIO) and FIFO buffers. The receive PIO employs a DMA ring buffer backup so incoming packets can be copied directly into host memory when the PIO FIFO buffer is full. The adapter may be programmed to generate early receive interrupts when only a portion of a packet has been received from the network, so as to decrease latency. The adapter may also be programmed to generate a second early interrupt so that the copying of a large

packet to the host may overlap reception of the packet end. The adapter to begin packet transmission before the packet is completely transferred from the host to the adapter, which further reduces latency. The minimal latency of the adapter allows it to employ receive and transmit FIFO buffers which are small enough to be contained within RAM internal to an Application Specific Integrated Circuit (ASIC) containing the transceiver, ethernet controller, FIFO control circuitry and the host interface as well.

EP 0 577 115 A3

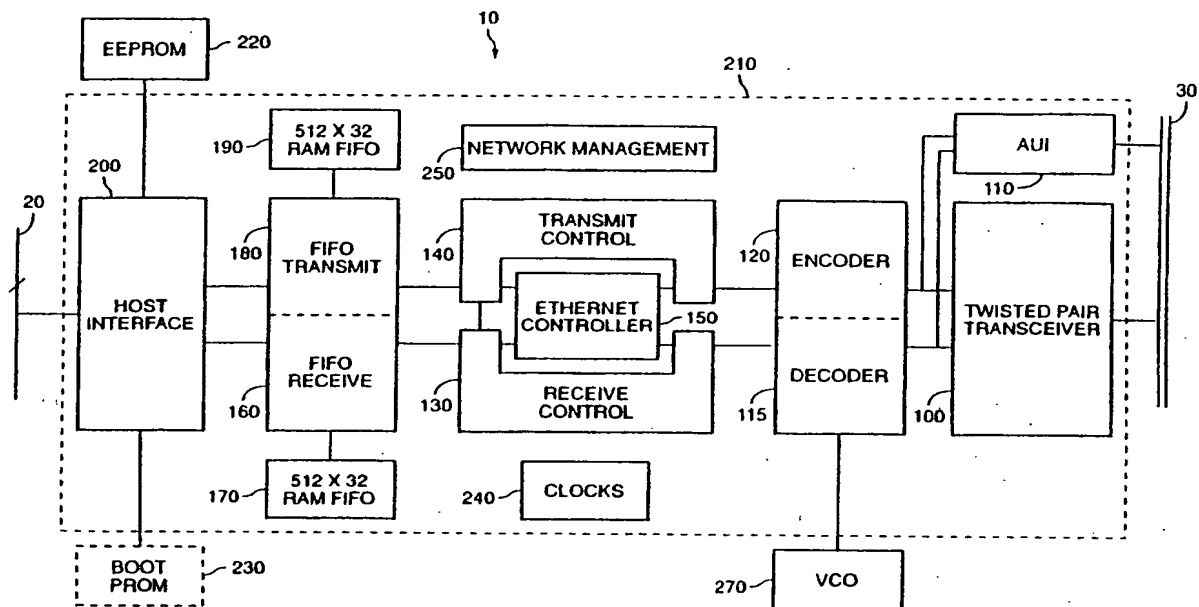


FIG. 1